10/089147
JC13 Rec'd PCT/PTO 27 MAR 2002

7/pRTS

The lipid body lipoxygenase ß-barrel

The present invention relates to a method of targeting proteins involved in lipid or fatty acid biosynthesis into liposomes or lipid bodies. The invention relates to a method of producing fatty acids and lipids in an oil-producing organism.

Moreover, the invention relates to nucleic acid sequences which encode a polypeptide and which is composed of a combination of nucleic acid sequences of a biosynthesis nucleic acid sequence of the fatty acid or lipid metabolism with a nucleic acid sequence which encodes the targeting of proteins. Moreover, the invention relates to nucleic acid constructs comprising these nucleic acid sequences, to vectors, and to transgenic organisms comprising the nucleic acids, nucleic acid constructs and/or vectors.

Oil crops play an important role in human nutrition. In germinating seeds, the endogenous stores of triacylgyceride-storing [sic] plants are degraded to make possible the *de novo* production of new tissue even in the absence of light [Rees et al., Biochim. Biophys. Acta, 385, 1975: 145-156; Kindl, H., Biochimie, 75, 1993: 225-230]. Triacylgycerides [sic] are stored in particularly highly specialized tissues, for example in the endosperm or the cotyledons. Cells of this type contain lipid bodies as compartments which store the fat [Murphy, D.J., Prog. Lipid Res., 32, 1993:247-280; Huang, A.H.C., Curr. Opinion Struct. Biol., 4, 1994: 493-498]. At the outset of germination, the lipid bodies and their contents are degraded and provide fatty acids for the glyoxysomes, which are responsible for fatty acid ß-oxidation [Kindl, H., Z. Naturforsch., C 52, 1997: 1-8]. In cucumber cotyledons, phospholipase A_2 (PLA) [May. C. et al., Biochim. Biophys. Acta, 1393, 1998: 267-276] and a specific isoform of lipoxygenase [lipid body lipoxygenase, LBLOX) [sic] [Feussner, I. et al., Planta, 198, 1998: 288-293; Höhne, M. et al., Eur. J. Biochem., 241, 1996: 6-11.] are synthesized during the step of triacylglyceride mobilization and transported to the lipid bodies. PLA (a patatin-like protein [May. C. et al., Biochim. Biophys. Acta, 1393, 1998: 267-276]) plays a decisive role in initiating the mobilization process by destroying the phospholipid monolayer of the lipid bodies [Noll, F. et al., J. Struct. Biol., 1999: submitted]. Lipoxygenase subsequently brings about modification of the acyl residues in the triacylgycerides [sic], which mainly take the form of linoleyl groups [Feussner, I. et al., FEBS Lett., 431, 1998: 433-436]. Following reduction and the effect of a specific hydroxyoctadecadienoyl-

dependent lipase, S-13-hydroxyoctadecadienoate is finally released from the lipid bodies into the cytosol [Feussner, I. et al., Proc. Natl. Acad. Sci. USA, 92, 1995: 11849-11853] and then broken down by the glyoxysomes [Kindl, H., Biochimie, 75, 1993: 225-230; Kindl, H., Z. Naturforsch., C 52, 1997: 1-8].

During lipid mobilization, a set of *de novo* proteins, among them LBLOX and PLA, is transported to the surface of the lipid bodies [Sturm, A. et al., Eur. J. Biochem., 150, 1985: 461-468]. Earlier work [May. C. et al., Biochim. Biophys. Acta, 1393, 1998: 267-276; Feussner, I. et al., Planta, 198, 1998: 288-293] has shown that LBLOX and PLA are synthesized transiently over a short period only, i.e. at the beginning of lipid mobilization. The processes during the transport of these proteins within the plant cell remain entirely unclear.

It is an object of the present invention to understand the signals of intracellular transport and to utilize them for the targeting of proteins.

We have found that this object is achieved by a method of targeting proteins involved in lipid or fatty acid biosynthesis into liposomes or lipid bodies, which comprises combining the protein-encoding nucleic acids and one of the following sequences to give a joint protein-encoding sequence:

a) a nucleic acid sequence with the sequence shown in SEQ ID NO: 1,

b) nucleic acid sequences which are derived from the nucleic acid sequence shown in SEQ ID NO: 1 as the result of the degeneracy of the genetic code,

c) derivatives of the nucleic acid sequence shown in SEQ ID NO: 1 which encode polypeptides with the amino acid sequences shown in SEQ ID NO: 2 and which have at least 60% homology at the amino acid level,

d) a nucleic acid sequence with the sequence shown in SEQ ID NO: 3 or the amino-terminal portion of the coding region of this sequence, and

and [sic] introducing the resulting sequence into a eukaryotic organism.

The present work has provided proof for the post-translational transfer of LBLOX and PLA toward lipid bodies. It has emerged that an N-terminal domain for the lipid body LOX, which is folded into a ß-barrel structure, is responsible for the membrane-binding properties of the enzyme. This domain of the protein encoded by the nucleic acid sequence according to the invention, which has a membrane-targeting structure, can be used in the method according to the invention for protein targeting [sic] foreign proteins toward lipid bodies, for example in oilseeds [sic].

The method according to the invention makes it possible to direct proteins which are advantageously involved in the fatty acid and/or lipid metabolism specifically to the site where synthesis is desired.

By introducing the nucleic acid sequence into a eukaryotic organism, proteins can be directed in the method according to the invention. To this end, the organisms are grown in a suitable medium.

The invention furthermore relates to a method of targeting proteins involved in lipid or fatty acid biosynthesis into liposomes or lipid bodies, which comprises introducing at least one nucleic acid sequence according to the invention as described hereinbelow or at least one nucleic acid construct into an oil-producing organism.

A method of producing fatty acids, which comprises introducing at least one nucleic acid sequence according to the invention or at least one nucleic acid construct into an oil-producing organism, growing this organism and isolating the oil contained in the organism.

A method of producing fatty acids, which comprises introducing at least one nucleic acid sequence according to the invention or at least one nucleic acid construct into an oil-producing organism, growing this organism, isolating the oil contained in the organism and liberating the fatty acids.

Method [sic] as described above, wherein the organism is a plant or a eukaryotic microorganism.

"Growing the organism" as described above is to be understood as referring both to the growing of plants and to the cultivation of eukaryotic microorganisms such as yeasts, fungi, ciliates, algae, or animal or plant cells or cell associations.

Organisms for the method which may be mentioned are, for example, plants such as Arabidopsis, barley, wheat, rye, oats, maize, soybean, rice, cotton, sugarbeet, tea, carrot, capsicum, canola, sunflower, flax, hemp, potato, triticale, tobacco, tomato, oilseed rape, coffee, tapioca, cassava, arrowroot, Tagetes, alfalfa, peanut, castor-oil plant, coconut, oil palm, safflower (Carthamus tinctorius), lettuce and the various tree, nut and grapevine species, or cocoa bean [sic], microorganisms such as yeasts like Yarrowia or Saccharomyces; fungi [lacuna] Mortierella, Saprolegnia, Traustochytrium or Pythium, algae or protozoa such as dinoflagellates such as Crypthecodinium. Preferred organisms are those which are naturally capable of synthesizing oils in substantial quantities, such as microorganisms like yeasts such as Yarrowia lypolytica [sic] or Saccharomyces cereviseae [sic], fungi such as Mortierella alpina, Pythium insidiosum or plants such as Arabidopsis thaliana, soybean, oilseed rape (Braccica [sic] napus), coconut, oil palm, canola, safflower (Carthamus tinctorius), castor-oil plant, Calendula, linseed (Linium [sic] usitatissimum), borage, peanut, cocoa bean [sic] or sunflower, with soybean, oilseed rape or sunflower being especially preferred.

The organisms obtained in the in the [sic] methods [sic] according to the invention advantageously contain saturated or unsaturated fatty acids in the form of bound fatty acids, i.e. the unsaturated fatty acids are predominantly in the form of the mono-, di- or triglycerides, glycolipids, lipoproteins or phospholipids such as oils or lipids, or else as fatty acids bound as esters or amides. Free fatty acids too are present in the organisms in the form of the free fatty acids or in the form of their salts. The organisms obtained in the method according to the invention by growing, and the saturated or unsaturated fatty acids which they contain, can be used directly for example for the production of pharmaceutical preparations, agrochemicals, feedstuffs or foods, or else first be isolated from the organisms. All of the isolation stages of the saturated or unsaturated fatty acids may be used, i.e. [lacuna] from fatty acid crude extracts up to fully purified fatty acids are suitable for making the above products. In an advantageous embodiment, the bound fatty acids can be liberated from, for example, the oils or lipids by, for example, basic hydrolysis, for example using NaOH or KOH. These free fatty acids can be used for the production of

pharmaceutical preparations, agrochemicals, feedstuffs or foods either directly in the mixture obtained or after further purification. The bound or free fatty acids may also be used for transesterification or esterification for example with other mono-, di- or triglycerides or glycerol in order to increase the proportion of unsaturated fatty acids in these compounds, for example in the triglycerides.

The organisms used in the methods are grown or cultured in the manner known to the skilled worker, depending on the host organism. Microorganisms such as bacteria, fungi, ciliates, or plant or animal cells are, as a rule, grown in a liquid medium comprising a carbon source, usually in the form of sugars, a nitrogen source, usually in the form of organic nitrogen sources such as yeast extract or salts such as ammonium sulfate, trace elements such as iron salts, manganese salts, magnesium salts and, if appropriate, vitamins, at temperatures of between 0°C and 100°C, preferably of between 10°C to [sic] 60°C, while passing in oxygen or in the absence of oxygen, depending on the organism. The pH of the liquid medium may be kept at a fixed value, i.e. the pH is controlled during the cultivation, or else the pH is not controlled and changes during the cultivation. Cultivation can be effected batchwise, semibatchwise or continuously. Nutrients can be provided at the outset of the fermentation or fed in semibatchwise or continuously. Cultivation on solid media is also possible.

After transformation, plants are, as a rule, first regenerated and then grown on or planted as usual. This can be done in the greenhouse or in the open.

After the organisms have been grown, the lipids can be obtained in the customary manner. To this end, the organisms can first be disrupted after harvesting or else be used directly. The lipids are advantageously extracted with suitable solvents such as apolar solvents such as hexane or ethanol, isopropanol or mixtures such as hexane/isopropanol, phenol/chloroform/isoamyl alcohol at temperatures of between 0°C to [sic] 80°C, preferably of between 20°C to [sic] 50°C. As a rule, the biomass is extracted with an excess of solvent, for example an excess of solvent to biomass of 1:4. The solvent is subsequently removed, for example via a distillation step. Extraction can also be effected with supercritical CO_2. After the extraction, the remainder of the biomass can be removed for example by filtration.

The crude oil thus obtained can subsequently be purified further, for example by removing suspended matter by treatment with polar solvents such as acetone or chloroform, followed by filtration or centrifugation. A further purification via chromatographic methods, distillation or crystallization is also possible.

To obtain the free fatty acids from the triglycerides, the latter are usually hydrolyzed as described above.

The invention furthermore relates to isolated nucleic acid sequences which encodes [sic] a polypeptide and which is [sic] composed of a combination of the nucleic acid sequences of a biosynthesis nucleic acid sequence of the fatty acid or lipid metabolism and one of the following nucleic acids:

a) a nucleic acid sequence with the sequence shown in SEQ ID NO: 1,

b) nucleic acid sequences which are derived from the nucleic acid sequence shown in SEQ ID NO: 1 as the result of the degeneracy of the genetic code,

c) derivatives of the nucleic acid sequence shown in SEQ ID NO: 1 which encode polypeptides with the amino acid sequences shown in SEQ ID NO: 2 and which have at least 60% homology at the amino acid level,

d) a nucleic acid sequence with the sequence shown in SEQ ID NO: 3 or the amino-terminal portion of the coding region of this sequence.

These nucleic acid sequences according to the invention make possible the targeting of proteins in the method according to the invention.

Derivative(s) are to be understood as meaning, for example, functional homologs of the proteins encoded by SEQ ID NO: 1 or of their biological activity, that is to say proteins which [lacuna] the same biological reactions as those controlled by SEQ ID NO: 1. These genes also make possible an advantageous targeting of proteins. Biological activity is understood as meaning directing proteins, advantageously proteins which are involved in the fatty acid and/or lipid metabolism, within the cell.

The nucleic acid sequence(s) used in the method according to the invention (for the purposes of the application, the singular is to comprise the plural and vice versa) or fragments thereof can

be used advantageously for isolating further genomic sequences via homology screening.

The derivatives mentioned can be isolated for example from other eukaryotic organisms such as fungi, yeasts or plants such as, specifically, mosses.

Derivatives or functional derivatives of the sequence stated in SEQ ID NO: 1 are furthermore understood as meaning, for example, allelic variants which have at least 60% homology, advantageously at least 70% homology, preferably at least 80% homology, especially preferably at least 85% homology, very especially preferably 90% homology at the deduced amino acid level. The homology was calculated over the entire amino acid range. The programs used were PileUp, BESTFIT, GAP, TRANSLATE and BACKTRANSLATE (= part of the program package UWGCG, Wisconsin Package, Version 10.0—UNIX, January 1999, Genetics Computer Group, Inc., Deverux et al., Nucleic. Acid Res., 12, 1984: 387—395) (J. Mol. Evolution., 25, 351-360, 1987, Higgins et al., CABIOS, 5 1989: 151—153). The amino acid sequence deduced from the abovementioned nucleic acids can be seen from sequence SEQ ID NO: 2. Homology is understood as meaning identity, that is to say the amino acid sequences have at least 60% identity. The sequences according to the invention have at least 50%, preferably at least 60%, especially preferably 70%, very especially preferably at least 80% homology at the nucleic acid level.

Allelic variants encompass in particular functional variants which can be obtained from the sequence shown in SEQ ID NO: 1 by deletion, insertion or substitution of nucleotides, the biological activity of the derived proteins which are synthesized being retained, i.e. these proteins are still capable of targeting proteins.

Starting from the DNA sequence described in SEQ ID NO: 1 or parts of these sequences [sic], such DNA sequences can be isolated from other eukaryotes such as, for example, those mentioned above, for example using customary hybridization methods or PCR technology. These DNA sequences hybridize with the abovementioned sequences under standard conditions. Short oligonucleotides, advantageously oligonucleotides having a length of 20 to 50 nucleotides, are advantageously used for the hybridization. However, longer fragments of the nucleic acids according to the invention or the complete sequences may also be used for the hybridization. Depending on the nucleic acid used, viz. oligonucleotide, longer fragment or complete sequence, or depending on the type of

nucleic acid used for the hybridization, viz. DNA or RNA, these standard conditions vary. Thus, for example, the melting temperatures for DNA:DNA hybrids are approximately 10°C lower than those of DNA:RNA hybrids of the same length.

Depending on the nucleic acid, standard conditions are understood as meaning, for example, temperatures of between 42 and 58°C in an aqueous buffer solution with a concentration of between 0.1 to [sic] 5 x SSC (1 x SSC = 0.15 M NaCl, 15 mM sodium citrate, pH 7.2) or additionally in the presence of 50% formamide such as, for example, 42°C in 5 x SSC, 50% formamide. The hybridization conditions for DNA:DNA hybrids are advantageously 0.1 x SSC and temperatures of between approximately 20°C to [sic] 45°C, preferably between approximately 30°C to [sic] 45°C. The hybridization conditions for DNA:RNA hybrids are advantageously 0.1 x SSC and temperatures between approximately 30°C to [sic] 55°C, preferably between approximately 45°C to [sic] 55°C. These temperatures stated for the hybridization are examples of melting temperatures calculated for a nucleic acid with a length of approximately 100 nucleotides and a G + C content of 50% in the absence of formamide. The experimental conditions for DNA hybridization are described in specialist textbooks of genetics such as, for example, Sambrook et al., "Molecular Cloning", Cold Spring Harbor Laboratory, 1989, and can be calculated using formulae known to the skilled worker, for example as a function of the length of the nucleic acids, the type of hybrid or the G + C content. The skilled worker can find more information on hybridization in the following textbooks: Ausubel et al. (eds), 1985, Current Protocols in Molecular Biology, John Wiley & Sons, New York; Hames and Higgins (eds), 1985, Nucleic Acids Hybridization: A Practical Approach, IRL Press at Oxford University Press, Oxford; Brown (ed), 1991, Essential Molecular Biology: A Practical Approach, IRL Press at Oxford University Press, Oxford.

Derivatives are furthermore understood as meaning homologs of the sequence SEQ ID No: 1, for example eukaryotic homologs, truncated sequences, single-stranded DNA of the coding and noncoding DNA sequence or RNA of the coding and noncoding DNA sequence.

Homologs of the sequence SEQ ID NO: 1 are furthermore understood as meaning derivatives such as, for example, promoter variants. These variants can be modified by one or more nucleotide substitutions, by insertion(s) and/or deletion(s) without, however, functionality or activity of the promoters being adversely affected. Moreover, the activity of the promoters can be increased by modifying the promoter sequence, or the promoters

in their entirety can be substituted by more effective promoters, also promoters from different species of organism.

Derivatives are also advantageously to be understood as meaning variants whose nucleotide sequence in the range -1 to -2 000 before the start codon have [sic] been modified in such a way that gene expression and/or protein expression is modified, preferably increased. Moreover, derivatives are also understood as meaning variants whose 3' end has been modified.

The nucleic acid sequences which encode the proteins according to the invention can be prepared synthetically or obtained naturally or comprise a mixture of synthetic and natural DNA constituents and also be composed of a variety of heterologous gene segments of a variety of organisms. In general, synthetic nucleotide sequences are prepared with codons which are preferred by the host organisms in question, for example plants. As a rule, this leads to optimal expression of the heterologous genes. These codons which are preferred by plants can be determined from codons with the highest protein frequency which are expressed in most of the plant species of interest. An example of Corynebacterium glutamicum is given in: Wada et al. (1992) *Nucleic Acids Res.* 20:2111-2118) [sic]. Such experiments can be carried out with the aid of standard methods and are known to the skilled specialist worker.

Functional equivalent sequences which encode the proteins according to the invention are those derivatives of the sequence according to the invention which retain the desired functions, that is to say the biological activity of the proteins, despite a deviating nucleotide sequence. Functional equivalents thus encompass naturally occurring variants of the sequences described herein, but also artificial artificial [sic] nucleotide sequences, for example those which have been obtained by chemical synthesis and which are adapted to the codon usage of a plant.

Also suitable are artificial DNA sequences as long as they confer the desired property, as described above, for example targeting in the fatty acid and/or lipid metabolism. Such artificial DNA sequences can be determined for example by backtranslating proteins which have been constructed by means of molecular modeling and which have the biological activity, or by in vitro selection. Techniques which are possible for the in vitro evolution of DNA for modifying or improving the DNA sequences have been described by Patten, P.A. et al., Current Opinion in Biotechnology 8, 724-733 (1997) or by Moore, J.C. et al., Journal of Molecular Biology 272, 336–347 (1997). Coding DNA sequences

which have been obtained by backtranslating a polypeptide sequence in accordance with the host-plant-specific codon usage are especially suitable.

Components of the nucleic acids according to the invention which may be mentioned are biosynthesis genes of the fatty acid and/or lipid metabolism such as advantageously a sequence which encodes proteins from among the following group of proteins:

Acyl-CoA dehydrogenase(s), Acyl-ACP [= acyl carrier protein] desaturase(s), Acyl-ACP thioesterase(s), fatty acid acyltransferase(s), fatty acid synthase(s), fatty acid hydroxylase(s), acetyl-coenzyme A carboxylase(s), acyl-coenzyme A oxidase(s), fatty acid desaturase(s), fatty acid acetylenases, lipoxygenases, triacylglycerol lipases, allenoxide synthases, hydroperoxide lyases and/or fatty acid elongase(s). They are preferably nucleic acids which encode one of the following proteins: fatty acid acyltransferase(s), Δ4 desaturase, Δ5 desaturase, Δ6 desaturase, Δ9 desaturase, Δ12 desaturase, Δ15 desaturase and/or a fatty acid elongase.

The nucleic acid sequences mentioned encode encode [sic] what are known as fusion proteins, component of the fusion protein being a polypeptide with the sequence mentioned in SEQ ID NO: 2 or a functionally equivalent part thereof. The second moiety of the fusion protein can be, for example, a further polypeptide with enzymatic activity, such as, for example, the abovementioned proteins.

In the method according to the invention, the nucleic acids can be combined advantageously with further genes of fatty acid biosynthesis. Examples of such genes are the acetyltransferases, further desaturases or elongases of unsaturated or saturated fatty acids as described in WO 00/12720. The combination with, for example, NADH cytochrome B5 reductases, which are capable of accepting or donating reduction equivalents, is advantageous for the in vivo and, specifically, the in vitro synthesis.

The proteins used in the method according to the invention are understood as meaning proteins which comprise an amino acid sequence shown in sequence SEQ ID NO: 2 or a sequence obtainable therefrom by substitution, inversion, insertion or deletion of one or more amino acid residues, the biological activity of the protein shown in SEQ ID NO: 2 being retained or not being reduced substantially. The term "not reduced substantially" is understood as meaning all those proteins which retain at least 10%, preferably 20%, especially preferably 30%, of the biological

activity of the original protein. For example, specific amino acids can be replaced by those with similar physicochemical properties (spatial dimension, basicity, hydrophobicity and the like) in this context. For example, arginine residues are exchanged for lysine residues, valine residues for isoleucine residues, or aspartic acid residues for glutamic acid residues. However, one or more amino acids may also be swapped round, added or removed, or several of these measured may be combined with each other.

Derivatives are also understood as meaning functional equivalents comprising in particular also natural or artificial mutations of an originally isolated sequence encoding the proteins according to the invention, these equivalents retaining the desired function, that is to say that their biological activity is not substantially reduced. Mutations comprise substitutions, additions, deletions, exchanges or insertions of one or more nucleotide residues. The present invention therefore also extends to those nucleotide sequences which are obtained by modification of the nucleotide sequence encoding the protein. The purpose of such a modification can be for example the further delimitation of the coding sequence comprised therein or else, for example, the insertion of further cleavage sites for restriction enzymes.

Functional equivalents are also those variants whose function is weakened (= not substantially reduced) or enhanced (= biological activity is greater than the activity of the original protein, i.e. activity is higher than 100%, preferably higher than 110%, especially preferably higher than 130%), compared with the original gene or gene fragment.

For introduction into a host organism, the nucleic acid sequences mentioned above which are used in the method according to the invention are advantageously inserted into an expression cassette (= nucleic acid construct). However, the nucleic acid sequences can also be introduced directly into the host organism. The nucleic acid sequence can advantageously be, for example, a DNA or cDNA sequence. Coding sequences which are suitable for insertion into an expression cassette are, for example, those which make possible the protein targeting according to the invention. These sequences can be of homologous or heterologous origin.

An expression cassette (= nucleic acid construct or nucleic acid fragment) is understood as meaning the sequence [lacuna] mentioned in SEQ ID NO: 1 which are the result of the genetic code and/or their functional or nonfunctional derivatives which

were linked operably to one or more regulatory signals, advantageously for increasing gene expression, and which control the expression of the coding sequence in the host cell. These regulatory sequences are intended to make possible the targeted expression of the genes and of protein expression [sic]. Depending on the host organism, this may mean, for example, that the gene is expressed and/or overexpressed only after induction, or that it is expressed and/or overexpressed immediately. For example, these regulatory sequences are sequences to which inductors or repressors bind, thus regulating the expression of the nucleic acid. In addition to these new regulatory sequences, or instead of these sequences, the natural regulation of these sequences before the actual structural genes may still be present and, if appropriate, may have been genetically modified so that the natural regulation has been eliminated and expression of the genes has been increased. However, the gene construct may also be simpler in construction, i.e. no additional regulatory signals were inserted before the nucleic acid sequence or its derivatives, and the natural promoter with its regulation was not removed. Instead, the natural regulatory sequence was mutated in such a way that regulation no longer takes place and/or gene expression is increased. These modified promoters can also be positioned by themselves in the form of part-sequences (= promoter together with parts of the nucleic acid sequences according to the invention) before the natural gene in order to increase the activity. Moreover, the gene construct may advantageously also comprise one or more of what are known as enhancer sequences linked operably to the promoter, and these enhancer sequences make possible an increased expression of the nucleic acid sequence. Additional advantageous sequences such as further regulatory elements or terminators may also be inserted at the 3' end of the DNA sequences. The nucleic acid sequences according to the invention may be present in the expression cassette (= gene construct) as one or more copies. Any further genes which, if appropriate, may be coexpressed and which are advantageously involved in fatty acid biosynthesis may be present in the expression cassette as one or more copies.

As described above, the regulatory sequences or factors can preferably exert a positive effect on, and thus increase, the gene expression of the genes which have been introduced. Thus, an enhancement of the regulatory elements may advantageously take place at the transcription level, by using strong transcription signals such as promoters and/or enhancers. In addition, however, increased translation is also possible, for example by improving the stability of the mRNA.

Promoters which are suitable as promoters in the nucleic acid construct are, in principle, all those which are capable of controlling the expression of foreign genes in organisms, advantageously in plants or fungi. In particular a plant promoter or plant promoters, for example those derived from a plant virus, are preferably used. Advantageous regulatory sequences for the method according to the invention are present for example in promoters such as cos, tac, trp, tet, trp-tet, lpp, lac, lpp-lac, lacI^q, T7, T5, T3, gal, trc, ara, SP6, λ-P_R or in the λ-P_L promoter, these promoters being used advantageously in Gram-negative bacteria. Further advantageous regulatory sequences are present, for example, in the Gram-positive promoters amy and SPO2, in the yeast or fungal promoters ADC1, MFα, AC, P-60, CYC1, GAPDH, TEF, rp28, ADH, or in the plant promoters such as CaMV/35S [sic] [Franck et al., Cell 21 (1980) 285-294], SSU, OCS, lib4, STLS1, B33, nos (= nopalin synthase promoter) or in the ubiquitin promoter. The expression cassette may also comprise a chemically inducible promoter by which the expression of the exogenous gene in the organisms, advantageously in the plants, can be controlled at a particular point in time. Such advantageous plant promoters are, for example, the PRP1 promoter [Ward et al., Plant. Mol. Biol. 22 (1993), 361-366], a benzenesulfonamide-inducible promoter (EP 388186), a tetracyclin inducible promoter (Gatz et al., (1992) Plant J. 2, 397-404), a salicylic acid-inducible promoter (WO 95/19443), an abscisic acid-inducible promoter (EP 335528), or an ethanol- or cyclohexanone-inducible promoter (WO 93/21334). Further plant promoters are, for example, the potato cytosolic FBPase promoter, the potato ST-LSI promoter (Stockhaus et al., EMBO J. 8 (1989) 2445-245 [sic]), the glycine max phosphoribosyl-pyrophosphate amidotransferase promoter (see also Genbank Accession Number U87999) or a node-specific promoter such as in EP 249676 can advantageously be used [sic]. Advantageous are, in particular, those plant promoters which ensure the expression in tissues or plant parts/organs in which the biosynthesis of fatty acids or their precursors takes place, such as, for example, in the endosperm or in the devloping embryo. Promoters which must be mentioned in particular are advantageous promoters which ensure the seed-specific expression such as, for example, the USP promoter or its derivatives, the LEB4 promoter, the phaseolin promoter or the napin promoter. The especially advantageous USP promoter or its derivatives, which are mentioned in accordance with the invention, mediate very early gene expression during seed development (Baeumlein et al., Mol Gen Genet, 1991, 225 (3): 459-67). Further advantageous seed-specific promoters which can be used for monocotyledonous and dicotyledonous plants are the promoters which are suitable for dicots such as, likewise being mentioned by way of example,

[sic] the oilseed rape napin gene promoter (US 5,608,152), the Arabidopsis oleosin promoter (WO 98/45461), the Phaseolus vulgaris phaseolin promoter (US 5,504,200), the Brassica Bce4 promoter (WO 91/13980) or the legume B4 promoter (LeB4, Baeumlein et al., Plant J., 2, 2, 1992: 233-239) or promoters which are suitable for monocots such as the promoters of the barley lpt2 or lpt1 gene (WO 95/15389 and WO 95/23230) or the promoters of the barley hordein gene, the rice glutelin gene, the rice oryzin gene, the rice prolamin gene, the wheat gliadin gene, the wheat glutelin gene, the maize zein gene, the oat glutelin gene, the sorghum kasirin gene or the rye secalin gene, which are described in WO 99/16890.

Furthermore preferred are, in particular, those promoters which ensure the expression in tissues or plant parts in which for example the biosynthesis of fatty acids, oils and lipids or their precursors takes place. Promoters which must be mentioned in particular are those which ensure seed-specific expression. Promoters which may be mentioned are the promoter of the oilseed rape napin gene (US 5,608,152), the Vicia faba USP promoter (USP = unknown seed protein, Baeumlein et al., Mol Gen Genet, 1991, 225 (3): 459-67), the promoter of the Arabidopsis oleosin gene (WO 98/45461), of the phaseolin promoter (US 5,504,200), or the promoter of the legumin B4 gene (LeB4; Baeumlein et al., 1992, Plant Journal, 2 (2): 233-9). Promoters which mediate seed-specific expression in monocotyledonous plants, such as the promoter of the barley lpt2 or lpt1 gene (WO 95/15389 and WO 95/23230), may furthermore be mentioned.

As described above, further genes to be introduced into the organism may additionally be present in the expression cassette (= gene construct, nucleic acid construct). These genes can be subject to separate regulation or else subject to the same regulatory region as the nucleic acids according to the invention. For example, these genes take the form of further biosynthesis genes, advantageously of fatty acid biosynthesis, which make possible an increased synthesis. Examples which may be mentioned are the genes for $\Delta 15$, $\Delta 12$, $\Delta 9$, $\Delta 6$, $\Delta 5$, $\Delta 4$ desaturase, the various hydroxylases, the acyl–ACP thioesterases, β–ketoacyl synthases or β–ketoacyl reductases. The desaturase genes are advantageously used in the nucleic acid construct.

In principle, all natural promoters together with their regulatory sequences, such as those mentioned above, can be used for the expression cassette according to the invention and for

the method according to the invention, as described hereinabove. In addition, synthetic promoters may also advantageously be used.

A variety of DNA fragments can be manipulated in order to obtain a nucleotide sequence which expediently reads in the correct direction and which is equipped with the correct reading frame. To connect the DNA fragments (= nucleic acids according to the invention) to each other, adaptors or linkers may be attached to the fragments.

The promoter and terminator regions can expediently be provided, in the direction of transcription, with a linker or polylinker containing one or more restriction sites for the insertion of this sequence. As a rule, the linker has 1 to 10, in most cases 1 to 8, preferably 2 to 6, restriction sites. In general, the linker within the regulatory regions has a size of less than 100 bp, frequently less than 60 bp, but at least 5 bp. The promoter can be both native, or homologous, and foreign, or heterologous, with regard to the host organism, for example the host plant. In the 5'-3' direction of transcription, the expression cassette comprises the promoter, a DNA sequence which encodes a gene [sic] used in the method according to the invention, and a region for transcriptional termination. Various termination regions can be exchanged for each other as desired.

Furthermore, manipulations which provide suitable restriction cleavage sites or which remove surplus DNA or restriction cleavage sites may be employed. Where insertions, deletions or substitutions such as, for example, transitions and transversions are suitable, *in vitro* mutagenesis, primer repair, restriction or ligation may be used. In the case of suitable manipulations such as, for example, restriction, chewing back or filling in of overhangs for blunt ends, complementary ends of the fragments may be used for ligation.

Attaching the specific ER retention signal SEKDEL (Schouten, A. et al., Plant Mol. Biol. 30 (1996), 781-792) may, inter alia, be of importance for an advantageous high expression level; the average expression level is tripled to quadrupled thereby. Other retention signals which occur naturally in plant and animal proteins located in the ER may also be employed for constructing the cassette.

Preferred polyadenylation signals are plant polyadenylation signals, preferably those which essentially correspond to T-DNA polyadenylation signals from Agrobacterium tumefaciens, in particular of gene 3 of the T-DNA (octopine synthase) of the Ti

plasmid pTiACH5 (Gielen et al., EMBO J.3 (1984), 835 et seq.) or suitable functional equivalents.

A nucleic acid construct is generated by fusing a suitable promoter to a suitable nucleic acid sequence according to the invention and a polyadenylation signal, using customary recombination and cloning techniques as are described, for example, in T. Maniatis, E.F. Fritsch and J. Sambrook, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989) and in T.J. Silhavy, M.L. Berman and L.W. Enquist, Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) and in Ausubel, F.M. et al., Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley-Interscience (1987).

The DNA sequence encoding a cucumber lipid body lipoxygenase comprises all those sequence features which are necessary to achieve a localization which is correct for the site of fatty acid, lipid or oil biosynthesis. No further targeting sequences per se are therefore necessary. However, such a localization may be desirable and advantageous and may therefore be artificially modified or increased so that such fusion constructs are also a preferred advantageous embodiment of the invention.

Especially preferred sequences are those which ensure targeting into plastids. Under certain circumstances, targeting into other compartments (reviewed by Kermode, Crit. Rev. Plant Sci. 15, 4 (1996), 285-423), for example into into [sic] the vacuole, the mitochondrion [sic], the endoplasmic reticulum (ER), peroxisomes, lipid bodies or, owing to the absence of suitable operable sequences, residence in the compartment of formation, the cytosol, may also be desirable.

The nucleic acid sequences encoding the proteins according to the invention, together with at least one reporter gene, are advantageously cloned into an expression cassette which is introduced into the organism via a vector or directly into the genome. This reporter gene should allow easy detectability via a growth, fluorescence, chemoluminescence, bioluminescence or resistance assay or via a photometric measurement. Examples of reporter genes which may be mentioned are genes for resistance to antibiotics or herbicides, hydrolase genes, fluorescence protein genes, bioluminescence genes, sugar or nucleotide metabolism genes, or biosynthesis genes such as the Ura3 gene, the Ilv2 gene, the luciferase gene, the β-galactosidase gene, the gfp gene, the 2-deoxyglucose—
6-phosphate phosphatase gene, the β-glucuronidase gene,

β-lactamase gene, the neomycin phosphotransferase gene, the hygromycin phosphotransferase gene, or the BASTA (= gluphosinate resistance) [sic] gene. These genes allow the transcription activity, and thus gene expression, to be measured and quantified readily. This makes possible the identification of sites in the genome which show different productivity.

In a preferred embodiment, an expression cassette comprises upstream, i.e. at the 5′ end of the coding sequence, a promoter and downstream, i.e. at the 3′ end, a polyadenylation signal and, if appropriate, further regulatory elements which are linked operably to the interposed coding sequence DNA sequence [sic]. Operable linkage is to be understood as meaning the sequential arrangement of promoter, coding sequence, terminator and, if appropriate, further regulatory elements in such a way that each of the regulatory elements can fulfill its intended function upon expression of the coding sequence. The sequences which are preferred for operable linkage are targeting sequences for ensuring subcellular localization.

An expression cassette may comprise, for example, a constitutive promoter (preferably the USP or napin promoter) and the gene to be expressed.

For expression in a prokaryotic or eukaryotic host organism, for example a microorganism such as a fungus, or a plant, the expression cassette is advantageously inserted into a vector such as, for example, a plasmid, a phage or other DNA which makes possible optimal expression of the genes in the host organism. Suitable plasmids are, for example, in E. coli pLG338, pACYC184, the pBR series such as, for example pBR322, the pUC series such as pUC18 or pUC19, the M113mp series, pKC30, pRep4, pHS1, pHS2, pPLc236, pMBL24, pLG200, pUR290, pIN-III^{113}-B1, λgt11 or pBdCI, in Streptomyces pIJ101, pIJ364, pIJ702 or pIJ361, in Bacillus pUB110, pC194 or pBD214, in Corynebacterium pSA77 or pAJ667, in fungi pALS1, pIL2 or pBB116; further advantageous fungal vectors are described by Romanos, M.A. et al., [(1992) "Foreign gene expression in yeast: a review", *Yeast* 8: 423-488] and by van den Hondel, C.A.M.J.J. et al. [(1991) "Heterologous gene expression in filamentous fungi] and in More Gene Manipulations in Fungi [J.W. Bennet & L.L. Lasure, eds., p. 396-428: Academic Press: San Diego] and in "Gene transfer systems and vector development for filamentous fungi" [van den Hondel, C.A.M.J.J. & Punt, P.J. (1991) in: Applied Molecular Genetics of Fungi, Peberdy, J.F. et al., eds., p. 1-28, Cambridge University Press: Cambridge]. Advantageous yeast vectors are, for example, 2μM, pAG-1, YEp6, YEp13 or pEMBLYe23. Examples of algal vectors or

plant vectors are pLGV23, pGHlac$^+$, pBIN19, pAK2004, pVKH or pDH51 (see Schmidt, R. and Willmitzer, L., 1988). The abovementioned vectors or derivatives of the abovementioned vectors constitute a small selection of the plasmids which are possible. Further plasmids are well known to the skilled worker and can be found, for example, in the book Cloning Vectors (Eds. Pouwels P.H. et al. Elsevier, Amsterdam-New York-Oxford, 1985, ISBN 0 444 904018). Suitable plant vectors are described, inter alia, in "Methods in Plant Molecular Biology and Biotechnology" (CRC Press), Chapter 6/7, pp. 71-119. Advantgeous vectors are what are known as shuttle vectors or binary vectors, which replicate in E. coli and Agrobacterium.

In addition to plasmids, vectors are also to be understood as meaning all of the other vectors known to the skilled worker such as, for example, phages, viruses such as SV40, CMV, baculovirus, adenovirus, transposons, IS elements, phasmids, phagemids, cosmids, linear DNA or circular DNA. These vectors are capable of replicating autonomously in the host organism or can be replicated chromosomally; chromosomal replication is preferred.

In a further embodiment of the vector, the expression cassette according to the invention may also advantageously be introduced into the organisms in the form of a linear DNA and integrated into the genome of the host organism via heterologous or homologous recombination. This linear DNA may be composed of a linearized plasmid or only of the expression cassette as vector, or the nucleic acid sequences according to the invention.

In a further advantageous embodiment, the nucleic acid sequence according to the invention may also be introduced into an organism by itself.

If, in addition to the nucleic acid sequence according to the invention, further genes are to be introduced into the organism, they may all be introduced into the organism together with a reporter gene in a single vector, or each individual gene with a reporter gene in one vector each, or several genes together in various vectors, it being possible to introduce the various vectors simultaneously or in succession.

The vector advantageously comprises at least one copy of the nucleic acid sequences according to the invention and/or of the expression cassette.

For example, the plant expression cassette can be incorporated into the transformation vector pRT ((a) Toepfer et al., 1993, Methods Enzymol., 217: 66-78; (b) Toepfer et al. 1987, Nucl. Acids. Res. 15: 5890 et seq.).

As an alternative, a recombinant vector (= expression vector) may also be transcribed and translated in vitro, for example by using the T7 promoter and T7 RNA polymerase.

Expression vectors used in prokaryotes frequently exploit inducible systems with and without fusion proteins or fusion oligopeptides, it being possible for these fusions to be effected at the N terminus or at the C terminus or other utilizable domains of a protein. In general, the purpose of such fusion vectors is: i.) to increase the expression rate of the RNA, ii.) to increase the achievable protein synthesis rate, iii.) to increase the solubility of the protein, or iv.) to simplify purification by means of a binding sequence which can be exploited in affinity chromatography. Also, proteolytic cleavage sites are frequently introduced via fusion proteins, which makes possible the elimination of a portion of the fusion protein also of [sic] purification. Such recognition sequences for proteases recognize [sic] are, for example, factor Xa, thrombin and enterokinase.

Typical advantageous fusion and expression vectors are pGEX [Pharmacia Biotech Inc; Smith, D.B. and Johnson, K.S. (1988) *Gene* 67: 31-40], pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ), which comprises glutathione S transferase (GST), maltose bind protein, or protein A.

Further examples of E. coli expression vectors are pTrc [Amann et al., (1988) *Gene* 69:301-315] and pET vectors [Studier et al., Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, California (1990) 60-89; Stratagene, Amsterdam, Netherlands].

Further advantageous vectors for use in yeast are pYepSec1 (Baldari, et al., (1987) *Embo J.* 6:229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz et al., (1987) *Gene* 54:113-123), and pYES derivatives (Invitrogen Corporation, San Diego, CA). Vectors for use in filamentous fungi are described in: van den Hondel, C.A.M.J.J. & Punt, P.J. (1991) "Gene transfer systems and vector development for filamentous fungi", in: Applied Molecular Genetics of Fungi, J.F. Peberdy, et al., eds., p. 1-28, Cambridge University Press: Cambridge.

As an alternative, insect cell expression vectors may also be used with advantage, for example for expression in Sf 9 cells. Examples of these are the vectors of the pAc series (Smith et al. (1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

Moreover, plant cells or algal cells may advantageously be used for gene expression. Examples of plant expression vectors are found in Becker, D., et al. (1992) "New plant binary vectors with selectable markers located proximal to the left border", *Plant Mol. Biol.* 20: 1195-1197 or in Bevan, M.W. (1984) "Binary *Agrobacterium* vectors for plant transformation", *Nucl. Acid. Res.* 12: 8711-8721.

Furthermore, the nucleic acid sequences according to the invention can be expressed in mammalian cells. Example [sic] of suitable expression vectors are pCDM8 and pMT2PC, which are mentioned in: Seed, B. (1987) *Nature* 329:840 or Kaufman et al. (1987) *EMBO J.* 6: 187-195). Promoters preferably to be used are of viral origin, such as, for example, promoters of polyoma[lacuna], adenovirus 2, cytomegalovirus or simian virus 40. Further prokaryotic and eukaryotic expression systems are mentioned in Chapters 16 and 17 in Sambrook et al., *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

In principle, the nucleic acids according to the invention, the expression cassette or the vector can be introduced into organisms, for example into plants, by all the methods which are familiar to the skilled worker.

For microorganisms, the skilled worker will find suitable methods in the textbooks by Sambrook, J. et al. (1989) Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, by F.M. Ausubel et al. (1994) Current protocols in molecular biology, John Wiley and Sons, by D.M. Glover et al., DNA Cloning Vol.1, (1995), IRL Press (ISBN 019-963476-9), by Kaiser et al. (1994) Methods in Yeast Genetics, Cold Spring Habor Laboratory Press or Guthrie et al. Guide to Yeast Genetics and Molecular Biology, Methods in Enzymology, 1994, Academic Press.

The transfer of foreign genes into the genome of the plant is termed transformation. It exploits the above-described methods of transforming and regenerating plants from plant tissues or plant cells for transient or stable transformation. Suitable methods are protoplast transformation by polyethylene glycol-induced DNA

uptake, the biolistic method with the gene gun - known as the particle bombardment method -, electroporation, incubation of dry embryos in DNA-containing solution, microinjection and Agrobacterium-mediated gene transfer. The methods mentioned are described, for example, in B. Jenes et al., Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, edited by S.D. Kung and R. Wu, Academic Press (1993) 128-143 and in Potrykus Annu. Rev. Plant Physiol. Plant Molec. Biol. 42 (1991) 205-225) [sic]. The construct to be expressed is preferably cloned into a vector which is suitable for transforming Agrobacterium tumefaciens, for example pBin19 (Bevan et al., Nucl. Acids Res. 12 (1984) 8711). Agrobactgeria transformed with such a vector can then be used for transforming plants, in particular crop plants such as, for example, tobacco plants, for example by bathing scarified leaves or leaf sections in an agrobacterial solution and subsequently growing them in suitable media. The transformation of plants with Agrobacterium tumefaciens is described, for example, by Höfgen and Willmitzer in Nucl. Acid Res. (1988) 16, 9877 or is known, inter alia, from F.F. White, Vectors for Gene Transfer in Higher Plants; in Transgenic Plants, Vol. 1, Engineering and Utilization, edited by S.D. Kung and R. Wu, Academic Press, 1993, pp. 15-38.

Agrobacteria transformed with an expression vector as described above can also be used in the known manner for transforming plants such as test plants such as Arabidopsis or crop plants such as cereals, maize, oats, rye, barley, wheat, soybean, rice, cotton, sugarbeet, canola, triticale, rice [sic], sunflower, flax, hemp, potato, tobacco, tomato, coffee, cacao, tea, carrot, capsicum, oilseed rape, tapioca, cassava, arrowroot, Tagetes, alfalfa, lettuce and the various tree, nut and grapevine species, in particular oil-containing crop plants such as soybean, peanut, castor-oil plant, borage, linseed, sunflower, canola, cotton, flax, oilseed rape, coconut, oil palm, safflower (Carthamus tinctorius) or cocoa bean [sic], for example by bathing scarified leaves or leaf sections in an agrobacterial solution and subsequently growing them in suitable media.

The genetically modified plant cells can be regenerated via all methods known to the skilled worker. Suitable methods can be found in the abovementioned publications by S.D. Kung and R. Wu, Potrykus or Höfgen and Willmitzer.

Suitable organisms or host organisms for the nucleic acids used [lacuna] the method according to the invention, the nucleic acid construct used or the vector used are, in principle, all those organisms which are capable of synthesizing fatty acids,

specifically unsaturated fatty acids, and which are suitable for the expression of recombinant genes. Examples which may be mentioned are plants such as Arabidopsis, Asteraceae such as Calendula or crop plants such as Brassica, Linium [sic], soybean, peanut, castor-oil plant, sunflower, maize, cotton, flax, oilseed rape, coconut, oil palm, safflower (Carthamus tinctorius) or cocoa bean [sic], microorganisms such as yeasts, for example the genera Yarrowia or Saccharomyces, fungi, for example the genus Mortierella, Saprolegnia or Pythium, bacteria such as the genus Escherichia, cyanobacteria, ciliates, algae or protozoans such as dinoflagellates, such as Crypthecodinium. Preferred organisms are those which are naturally capable of synthesizing oils in substantial amounts, like yeasts such as Saccharomyces cereviseae [sic] or Yarrowia lypolytica [sic], fungi of the genera Mortierella, Traustochytrium or Pythium such as Mortierella alpina, Pythium insidiosum, or plants such as Brassica napus, Linium [sic] usitatissimum, soybean, oilseed rape, coconut, oil palm, safflower, castor-oil plant, Calendula, peanut, cocoa bean [sic] or sunflower, with soybean, oilseed rape, sunflower, castor-oil plant, Mortierella or Pythium being especially preferred. In principle, transgenic animals, for example C. elegans, are also suitable as host organisms.

Other useful host cells are mentioned in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990).

Expression strains which can be used, for example those which have a lower protease activity, are described in: Gottesman, S., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128.

Depending on the choice of the promoter, the expression of the nucleic acid sequences according to the invention can take place specifically in the leaves, in the seeds, in the tubers or other parts of the plant. The present invention furthermore relates to transgenic plants overproducing such fatty acids, oils or lipids, their propagation material and their plant cells, plant tissue or plant parts. The invention preferably relates to transgenic plants, for example crop plants such as maize, oats, rye, wheat, barley, maize, rice, soybean, sugarbeet, canola, triticale, sunflower, flax, hemp, tobacco, tomato, coffee, cacao, tea, carrot, capsicum, oilseed rape, tapioca, cassava, arrowroot, Tagetes, alfalfa, lettuce and the various tree, nut and grapevine species, potato, in particular oil-containing crop plants such as soybean, peanut, castor-oil plant, borage, linseed, sunflower, canola, cotton, flax, oilseed rape, coconut, oil palm, safflower

(Carthamus tinctorius) or cocoa bean [sic], test plants such as Arabidopsis or other plants such as mosses or algae comprising a functional nucleic acid sequence according to the invention or a functional expression cassette. Functional in this context means a biologically active protein is formed.

The expression cassette or the nucleic acid sequences comprising a nucleic acid sequence according to the invention can additionally also be employed for transforming the organisms mentioned above by way of example, such as bacteria, cyanobacteria, filamentous fungi, ciliates, animals or algae, with the purpose of increasing the fatty acid, oil or lipid content. Preferred transgenic organisms are yeasts, fungi or plants, especially preferably plants.

Transgenic organisms are understood as meaning organisms which comprise a foreign nucleic acid which originates from another organism and encodes a protein used in the method according to the invention. The term transgenic organisms is also understood as meaning organisms comprising a nucleic acid from the same organism, with nucleic acid being present as an additional gene copy or not in the natural nucleic acid environment of the nucleic acid sequence according to the invention. Transgenic organisms are also organisms in which the natural 3' and/or 5' region of the nucleic acid according to the invention has been modified over the starting organism by targeted recombinant modifications. Preferred transgenic organisms are those into which a foreign DNA has been introduced. Transgenic plants are especially preferred. Transgenic plants are understood as meaning individual plant cells and their cultures such as, for example, callus cultures on solid or in liquid media, as well as plant parts and intact plants.

Examples:

Plant material and cell fractionation

Cucumber seeds were germinated in the dark at 26°C for 2 or 4 days, as stated for each individual preparation. The cotyledons were harvested and homogenized by cutting them up with a surgical blade as previously described [Kindl, H. et al., Methods Enzymol., 96, 1983: 700-715]. After cell debris has been removed and the material subjected to differential centrifugation, the sediments of the centrifugation at 10 000 x g or the centriguation at 100 000 x g were sedimented or floated in a sucrose gradient as described by [Sturm, A. et al., FEBS Lett., 160, 1983: 165-168]. A crude lipid body fraction was obtained by

subjecting the supernatant of a brief (10 minutes) centrifugation at 2 000 x g to centrifugation for 30 minutes at 10 000 x g and removing the lipid layer formed in the upper part of the centrifuge tube. The lipid bodies were further purified by gently suspending the lipid layer and repeated flotation [Sturm, A. et al., Eur. J. Biochem., 150, 1985: 461-468].

Tobacco plants were grown in Magenta boxes at 22°C in permanent light (2 000 lux). The media for growing transformants were as described by [Horsch, R. B. et al., Science 227, 1985: 1229-1231; Jefferson, R. A. et al., EMBO J., 6, 1987: 3901-3907].

Plasmid constructions and protein preparation

CSLBLOX-221 in the vector pSport-1 [Höhne, M. et al., Eur. J. Biochem., 241, 1996: 6-11] was used for generating LBLOX deletion forms. An N-terminal deletion was prepared by cleaving CSLBLOX-221 first with *Sma*I/*Nde*I and subsequently with *Hae*III. After ligation into pSport-1 (Life Technologies), this construct, in which the first 80 nucleotides of pCSLBLOX-221 had been deleted [Höhne, M. et al., Eur. J. Biochem., 241, 1996: 6-11], was termed LBLOXΔ80N. After mRNA had been generated by means of *in vitro* transcription, translation of the corresponding protein started at a methionine residue, which corresponded to nucleotide 192 of pCSLBLOX-221, and the protein thus lacked the 48 N-terminal amino acid residues of the wild-type LBLOX. Both LBLOXΔ51 and LBLOXΔ96 were deletions in the C-terminal region downstream of amino acid residue 696. LBLOXΔ51 and LBLOXΔ96 had deletions of 51 and 96 amino acid residues, respectively, behind position 696, but retained the original C terminus. For these preparations, the *Msc*I site at nucleotide 2128 of p-CSLBLOX-221 was employed for the cleavage, and the upstream fragment was ligated to the corresponding C-terminal fragments generated by means of PCR. To construct LBLOXΔ504, the *Nde*I site of pCSLBLOX-221 at nucleotide 1450 and the *Aat*II site downstream of the translation stop of pCSLBLOX-221 were used to delete the entire C-terminal portion of LBLOX. After these sites had been made blunt-ended and religated, the corresponding protein, which was generated by in vitro transcription/translation of the construct, lacked the C-terminal half of the LBLOX molecule.

To clone a fusion protein GST-LBLOX244, a fragment derived from the 26 kDa glutathione-S transferase domain [Smith, D.B. et al., Gene, 67, 1988: 31-40] of *Schistosoma japonicum* was used as N-terminal portion and the N-terminal segment of LBLOX, which is composed of 244 amino acid residues, was used as C-terminal portion. Using the *Bam*HI/*Xho*I-cleaved gene fusion vector

pGEX-4T-3 (Pharmacia) and a fragment 732 nucleotides in length, which corresponded to amino acid residues 1-244 of LBLOX, for cloning and as affinity label for protein purification, the fusion protein GST-LBLOX244 was isolated by chromatography on glutathione-Sepharose 4B (Pharmacia).

The expression of LBLOX in bacteria was carried out after cloning the pCSLBLOX-221 insert into a pQE vector (Quiagen [sic]). To generate LBLOX by means of *in vitro* translation, the LBLOX sequence present in vector pSport-1, namely pCSLBLOX-221 [Höhne, M. et al., Eur. J. Biochem., 241, 1996: 6-11], was cleaved with *Aat*II, transcribed using T7 polymerase and translated in reticulocyte lysate. Equally, the patatin-like protein (which is identical with phospholipase A_2) was obtained by *in vitro* transcription/translation using pCSPAT-291 and the PLA cDNA under the control of the T7 promoter [May. C.et al., Biochim. Biophys. Acta 1393, 1998: 267-276].

Transfection of tobacco

Using the vector pBI121 (Clontech), which contains the cauliflower mosaic virus 35S RNA promoter, the coding region of the ß-glucuronidase gene and the NOS terminator, construct pBI121ΔGUS was generated by excising the GUS cassette with *Sma*I and *Sst*I and making it blunt-ended using T4 DNA polymerase. LBLOX present in the vector pSport-1 [Höhne, M. et al., Eur. J. Biochem., 241, 1996: 6-11] was excised with *Sma*I/*Bam*HI, ligated to a *Bam*HI linker, and introduced into the *Bam*HI-cleaved dephosphorylated vector pBI121ΔGUS. The product, pBI121ΔGUS-LBLOX, was either used for transformation into *Agrobacterium tumefaciens* or modified further to give pBI121ΔGUS-LBLOX-HA_3. This last-mentioned plasmid was constructed by using a single *Sac*I site behind nucleotide 252 of pCSLBLOX-221. Following cleavage with *Sac*I and dephosphorylation, a triple hemagglutinin label (HA label) provided with an *Sac*I site was inserted. Taking into consideration the 3 -YPYDVPDYA- sequences and the linkers, the overall length of the insert was 30 amino acids.

Agrobacterium tumefaciens LB-A4404 was transformed with pBI121ΔGUS-LBLOX or pBI121ΔGUS-LBLOX-HA_3 using the freeze-thaw method. These bacteria were used for transforming *Nicotiana tabacum* cv. Petit Havanna SR-1 leaf disks using the established method by Horsch et al. [Science 227, 1985: 1229-1231]. The shoots were selected on Linsmaier and Skoog medium supplemented with 0.5 mg/l N-benzylaminopurine, 500 mg/l cefotaxime and

75 mg/l kanamycin. Kanamycin-resistant plants with increased LOX contents were subjected to vegetative propagation.

Preparation of HA-labeled LBLOX by expression into tobacco

Leaves of homozygous kanamycin-resistant tobacco lines comprising the pBI121ΔGUS-LBLOX-HA_3 construct were homogenized in 50 mM Hepes-NaOH, pH 7.4, 5 mM $MgCl_2$, 0.5 mM EDTA, 50 mM KCl, 2.5 mM dithiothreitol, 2 mM phenylmethylsulfonyl fluoride and 15% (w/w) sucrose (buffer A) in the presence of polyvinylpyrrolidone (25 000) (Merck). After the centrifugations, the supernatant of the 100 000 x g centrifugation was demineralized, concentrated and fractionated on a large Biogel A-1.5 column. Fractions which corresponded to the 100 kDa range and which had been analyzed by Western blot using anti-HA antiserum were collected. This mutated LBLOX protein contained an insertion of 30 amino acid residues, which corresponded to a triple HA label, behind amino acid residue 68 of the wild-type enzyme and had a theoretical molecular weight of 104 kDa. This difference in size between the wild-type protein and the recombinant protein was detected unambiguously by means of SDS-PAGE.

Radiolabeling proteins by *in vitro* synthesis

Translation was effected using purified mRNAs, reticulocyte lysate and [^{35}S]L-methionine in the presence or absence of canine pancreas microsomes. As an alternative, microsomes prepared from cucumber cotyledons were used for cotranslational or posttranslational transport assays.

Labeling of proteins by *in vivo* protein synthesis

In vivo radiolabeling experiments were carried out as short-pulse experiments with cotyledons of seedlings which had germinated for 4 days. Five g of cotyledons were cut into 2 mm sections and incubated for 15 minutes with 8 MBq [^{35}S]L-methionine (40 TBq/mmol). Careful homogenization and preparation of subfractions was carried out as described above.

$^{45}Ca^{2+}$ feeding, and analysis of the lipid body fraction

Two g of cotyledons which had been collected from cucumber seedlings grown for 2.5 days (= d) in the dark at 26°C were cut into sections and incubated for 3 hours in the dark with 6 MBq $^{45}Ca^{2+}$ (8 GBq/mmol). After homogenization by comminuting using a surgical blade in the presence of buffer A, the homogenate was centrifuged for 25 minutes at 2 000 x g. After

removal of the top phase, which comprised the lipid bodies, and decanting off from the sediment, the extract was centrifuged for 1 hour at 100 000 x g.

The fraction which comprised the lipid bodies was resuspended in buffer A, brought to 30% (w/w) sucrose, and covered with a layer of buffer A. After centrifugation at 100 000 g for 1 hour, the floating lipid bodies were collected and subjected to a variety of wash methods.

Liposome preparation

Liposomes were prepared as described by [Woodle, M C. & Papahadjopoulos, D., Methods Enzymol. 171, 1989: 193-217] from a crude soybean lecithin mixture (Sigma) or from defined dilinoleoylphosphatidylcholine, with or without addition of the serine derivative, as unilamellar vesicles. Using detergent solubilization and removal of the former, the molar ratio of dilinoleoylphosphatidylcholine to sodium cholate was routinely 0.6:1.0. The efficacy at which the detergent had been removed by dialysis in an Amicon chamber equipped with PM10 membrane was monitored [Yamazaki, N. et al., Methods Enzymol. 242, 1994: 56-65]. Vesicle size was determined under the microscope by comparison with the size (10 $\pm$ 0.5 µm) of MonoQ beads (Pharmacia). For specific experiments, triacylglycerides (trilinolein) was incorporated into the liposomes to give rise to phospholipid-coated lipid droplets which are comparable to "black" lipid bodies. This preparation was started by dissolving 200 mg of soybean lecithin and 500 mg of trilinolein in 10 ml of methanol/chloroform (1:1). After removal of the solvent in vacuo, the residue was resuspended in dialysis buffer as described by [Yamazaki, N. et al., Methods Enzymol. 242, 1994: 56-65]. In each case, the vesicles obtained by flotation were analyzed by TLC. The liposomes had a diameter in the region of 1 µm.

Binding experiments in combination with the flotation assay

A liposome suspension corresponding to 1 mg of lecithin in 150 mM NaCl, 50 mM Tris-HCl, pH 8.0, was incubated for 10 minutes either with 4 µg of unlabeled protein or with the supernatant of a reticulocyte lysate translation mixture. After the suspension was brought to 42% (w/w) sucrose, it was transferred into a 12 ml centrifuge tube. The sample was covered with a layer of a linear sucrose density gradient of 37-26% (w/w) sucrose. The protein-covered liposomes were floated by centrifugation for

6 hours at 100 000 g. Fractionation was followed by protein analysis by means of SDS-PAGE and immunolabeling.

Binding experiments with lipid bodies or microsomes were carried out analagously. All densities in the sucrose gradients are stated in correlation with the sucrose concentration (w/w).

Immunological methods

The antisera against LBLOX [Sturm, A. et al., Eur. J. Biochem. 150, 1985: 461-468], PLA [May. C. et al., Biochim. Biophys. Acta 1393, 1998: 267-276] and isocitrate lyase [Frevert, J. & Kindl, H. Eur. J. Biochem., 92, 1978: 35-43] were produced in rabbits. Moreover, monoclonal antibodies against GST (Pharmacia) and the epitope of the hemagglutinin protein from human influenza virus (Boehringer) were used. Immunoprecipitations of the radiolabeled enzymes were carried out under standard conditions (method 1) by adding 1 µg of the corresponding purified protein to the mixture prior to precipitation with 20 µl of antiserum. The reaction mixture was left to stand for 12 hours at 20°C and for 20 hours at 4°C, and the precipitate was sedimented by centrifugation at 3 000 x g. The sediment was washed at least 5 times and then dissolved in SDS. For the direct precipitation of other proteins (method 2), the antigen was not diluted any further but was incubated for 6 hours with 2 µl of antiserum and subsequently mixed with protein A Sepharose. After the mixture had been transferred into a small vessel and washed thoroughly, the antigen together with the IgG was eluted using 100 mM acetic acid.

Further assays

To compare the protein structures, the gel was subjected to restricted proteolysis [Höhne, M. et al., Eur. J. Biochem. 241, 1996: 6-11]. The lipid was subjected to TLC analysis on Silica-Gel G (Merck) using methanol/chloroform/water 65:25:4 as the solvent system.

RESULTS

To study the sequence of steps required for transferring LBLOX from the ribosome to its final cellular target, studies first focused on which pools in the cell are crossed by the LBLOX protein. Secondly, the targeting structure, targeting sequence or targeting domains responsible for the transfer were localized.

In vivo and *in vitro* studies with LBLOX to differentiate between cotranslational transport into the ER and posttranslational transport to lipid bodies

No differences in the molecular weight of LBLOX translatated *in vitro* and the cellular forms bound to ER or lipid body membranes were observed (fig. 1A). This demonstrates that the transport to the membranes takes place without discernible chemical modification. Probable mechanisms by means of which LBLOX can reach its final location in the cell are both cotranslational transport to the ER, followed by transfer into the lipid bodies, and posttranslational transport to the lipid bodies via a cytosolic pool. Firstly, an *in vitro* translation of LBLOX mRNA was carried out under cotranslational transport conditions in the presence of canine pancreatic microsomes or cucumber cotyledon microsomes. Secondly, a similar protocol was used to test whether radiolabeled LBLOX translation product was also transported to the membranes when translation and addition of microsomes were effected one after the other (fig. 1B). In the latter case, the ribosomes were removed by centrifugation from de novo synthesized protein before microsomes were added. After reisolation of ER-enriched membranes by flotation, only part of the translated protein remained at the position where the suspension had been introduced onto the gradient (fig. No. 12). Some LBLOX migrated to densities corresponding to 33% sucrose (fig. No. 11) or 30% sucrose (fig. No. 9). Small amounts of LBLOX were found at the position of the smooth ER (fig. No.5-7).

The comparison of the amount of LBLOX radioactivity in the membranes which, in a series of experiments, had been reisolated by means of gradient centrifugation illustrated that the use of the arrangement for the post-translational transport to microsomes revealed identical or slightly higher amounts of membrane-bound LBLOX than the use of the protocol for cotranslational transport (data not shown). The detailed comparison of the migration behavior of the LBLOX primary translation product and LBLOX reisolated from the membranes after flotation revealed no discernible differences in molecular weight.

PLA-mRNA which had been obtained by means of *in vitro* transcription of pPAT291 was also translated, and the translation product, i.e. radiolabeled phospholipase, was incubated with microsomes. After flotation in a sucrose gradient, the subfractions were analyzed by means of SDS-PAGE and fluorography

(fig. 1C). Fluorography revealed that much of the translation product had bound to microsome membranes.

The results obtained by studying the *in vitro* transfer of LBLOX to microsomes (fig. 1B) were confirmed by the analysis of the *in vivo* situation by monitoring the sequence of the events taking place when LBLOX is transferred to lipid bodies. By using pulse-chase experiments, we determined the pools crossed by the de novo synthesized LBLOX on its route to the lipid bodies. This was done to find out whether the pool in which the labeled LBLOX was first detectable was the ER membrane fraction or the cytosol fraction.

Fig. 2A (lane 2) shows that most of the pulse-labeled LBLOX was obtained in the fraction comprising the cytosol. A small, but significant amount of the radioactive LBLOX was continuously found in the microsomes and also in microsome fractions which were purified further by flotation in sucrose gradients (fig. 2A). To rule out the possibility that small percentages of all de novo synthesized proteins contaminate the ER-containing fraction, we analyzed, in the microsome fraction, the amount of proteins which were either cytosolic or had been released as artefacts into the cytosol-containing supernatant of 100 000 g centrifugation. Using isocitrate lyase as strongly expressed cucumber protein at this developmental level, the presence of this protein in the microsome fraction and in the soluble supernatant was determined by immunoprecipitation. Fig. 2B shows that, in contrast to LBLOX, isocitrate lyase was virtually absent in the ER preparation.

The data shown in fig. 2A show that LBLOX crosses the cytosol as the first pool on its route toward the lipid bodies. It is also obvious that LBLOX has membrane-binding properties and is partially protected from proteolysis when bound to the ER membrane. Despite regions in the LBLOX molecule which have affinity for membranes, a major part of LBLOX is accessible to chemical modification *in vitro* and might therefore project into the cytosol *in vivo*. Part C of fig. 2 indicates the extent to which the part of LBLOX which is bound to microsome membranes is accessible to proteolytic degradation.

According to our findings, the LOX isoform which can be detected by Western blot analysis in microsomes isolated from cotyledons is identical to the LOX isoform bound to lipid bodies. This was demonstrated by comparing the fragment pattern following restricted proteolysis (data not shown). The way in which LBLOX is bound to the microsome membrane corresponds to that of a

peripherally bound membrane protein. Washing with 100 mM $MgCl_2$ removed over 90% of the labeled LBLOX. However, it should be emphasized that the binding of LBLOX to the microsomes in the presence of a low-salt buffer and also under the conditions of repeated gradient centrifugation was quite stable. The fact that binding was stable was also demonstrated by the observation that LOX on microsomes was only partially accessible to proteolysis (fig. 2, part C).

LBLOX's capacity of binding to microsome membranes was also demonstrated under a further *in vivo* condition. In a heterologous system, i.e. green tobacco leaves which expressed cucumber cotyledon LBLOX, the high expression of LBLOX resulted in considerable amounts of LBLOX binding to microsomes. Subfractionation of the 100 000 x g sediment, making use of a sucrose density gradient flotation, demonstrated that virtually all of the LBLOX previously sedimented by centrifugation at 100 000 x g remained bound to the floated membranes. The peak of the membrane-bound LBLOX (fig. 2, part D) agreed with the marker proteins for ER membranes.

To extend our chain of proof for direct posttranslational transfer of LBLOX to the membranes of lipid bodies, we studied the affinity of soluble LBLOX for microsome membranes *in vitro*. In a mixing experiment, we added a mixture of radioactive cytosolic proteins which had been isolated from cotyledons following pulse labeling (preparation A) to a supernatant of an extract which had been prepared from unlabeled cotyledons and centrifuged at 2 000 x g (preparation B). This procedure should allow the radioactive LBLOX, which was present as precursor in the cytosolic and membrane-free preparation (preparation A), to bind posttranslationally to microsome membranes, lipid bodies and other membranes not present in the nonradioactive extract (preparation B). Following brief incubation, we isolated microsomes as potential acceptor membranes for radioactive LBLOX (fig. 3A) and glyoxysomes as control (fig. 3B). Electrophoretic analysis demonstrated that the cytosolic LBLOX was bound to microsomes even under these conditions, which were similar to the *in vivo* situation. LBLOX was not detectible in the glyoxysome fraction, which was subjected to final purification by sucrose gradient flotation. Thus, soluble LBLOX binds to membranes, but not uniformly to all membranes. For example, the affinity of LBLOX for glyoxysome membranes (fig. 3B) is lower by at least one order of magnitude than the detected affinity for microsomes (fig. 3A, lane 3). Binding of LBLOX in cucumber cotyledons therefore requires a high degree of selectivity since other

organelles are not labeled with LBLOX and not modified for degradation.

Cucumber LBLOX and soybean LOX-1 differ with regard to their affinity for liposomes

To test whether LBLOX independently of specific protein-protein interactions has an intrinsic affinity for membrane lipids, the binding studies were extended and included liposomes as acceptor membranes. The experiments were first carried out with crude soybean lecithin as phosphatidylcholine source. Then, phosphatidylcholine with different degrees of purity was also used in combination with phosphatidylserine for generating liposomes. Liposome size was checked by comparison with MonoQ beads as standard.

The results of the liposome experiment (fig. 4) demonstrate the pronounced difference between the membrane affinity of cucumber LBLOX and that of soybean LOX-1. While the two proteins previously assigned to the cucumber lipid bodies, namely LBLOX and PLA, were almost quantitatively bound to the liposomes, the cytosolic LOX forms of cucumber and soybean LOX-1 had no affinity for the lipid phase. In control experiments with either a typical cytosolic protein or with a mixture which comprised LBLOX and cytosolic proteins and which had been prepared from cucumber cotyledons, it was possible to demonstrate that the membrane affinity of LBLOX is quite unique (data not shown).

The N-terminal region including the ß-barrel structure is essential for the binding of LBLOX to lipid bodies.

Taking into consideration the lacking proteolytic processing and the data obtained in the binding studies, it can be hypothesized that domains of a folded protein are responsible for the transfer to the membranes. Such a domain became likely when the amino acid sequence of LBLOX was matched with a structure based on the crystal structure data obtained for soybean LOX-1 at a resolution of 1.4 Å [Frevert, J. & Kindl, H., Eur. J. Biochem. 92, 1998: 35-43]. In contrast to soybean LOX-1, cucumber LBLOX showed not only a ß-barrel structure shortly downstream of the N-terminal section, but also had, in this domain, a quite unique arrangement of glutamyl residues which might be used for the coordination of Ca^{2+} (fig. 5). This, in turn, might suggest a relationship between the LBLOX ß-barrel and the members of Ca^{2+}-dependent membrane-binding domains.

To study the concept of a single domain as protein targeting means, a cDNA which encoded a fusion protein of glutathione-S transferase and the N-terminal LBLOX ß-barrel as C-terminal segment, was constructed. Following expression in bacteria, a 51 kDa protein which contained the 220 amino acid residues from glutathione-S transferase (and constitute the glutathione binding site) and the 224 N-terminal amino acid residues of LBLOX was isolated.

1 µg of the GST-LBLOX244 fusion protein and a lipid body suspension corresponding to 20 µg of protein were employed, and we observed an almost quantitative transfer of the fusion protein to the lipid bodies in addition to the LBLOX already present in the lipid bodies. The high extent to which the ß-barrel fusion protein was bound to the surface of the lipid bodies demonstrates that either the fusion protein competed well with the LBLOX originally bound to the lipid bodies or that the surface of the lipid bodies was not fully saturated with LBLOX. Fig. 6 compiles the indications that the N-terminal LBLOX ß-barrel alone suffices for a very effective binding to lipid bodies. Lane 5 shows the uptake of the 51 kDa fusion protein by the lipid bodies.

Two cDNA constructs, LBLOXΔ51 and LBLOXΔ96, were generated, and, following in vitro transcription/translation, the respective proteins were studied for binding to isolated lipid bodies. Both recombinant proteins were transported efficiently to lipid bodies (data not shown). These experiments were designed because the analysis of hydropathy plots of the LBLOX amino acid sequence suggested that a segment of amino acid residues around position 710 contains a hydrophobic region which might suffice for potential membrane binding. However, the binding assays were positive and demonstrated that the absence of the region around 710, as had been generated in the constructs LBLOXΔ51 and LBLOXΔ96, does not significantly reduce the efficiency of the binding of LBLOX to lipid bodies. Using the *in vitro* assays shown in fig. 1B, we also found that radiolabeled translation product was transported from LBLOXΔ504 to isolated lipid bodies. LBLOXΔ504 was only composed of the N-terminal half of the LBLOX molecule. This additionally confirms the concept that the C-terminal portion of LBLOX, which comprises the active center, is not necessary for the targeting. LBLOXΔ80N without the N-terminal extension, but with the ß-barrel, was transported to lipid bodies with a significantly lower efficiency than LBLOXΔ504.

By extending this type of *in vitro* experiments, it was found that the transfer to the lipid bodies is Ca^{2+}-independent (data not shown). Despite this result, the possibility that the Ca^{2+}-dependent binding of proteins to the surface of the lipid bodies might finally lead to the formation of a Ca^{2+}-enriched protein layer surrounding the lipid bodies was studied. To study the uptake of Ca^{2+} together with proteins such as LBLOX and PLA, we studied the formation of Ca^{2+}-covered lipid bodies by feeding $^{45}CaCl_2$ to cotyledons and subsequently isolating cell structures. After refloating, washing and treatment with 100 mM Na_2CO_3, the lipid body fraction contained 50 kBq of $^{45}Ca^{2+}$, corresponding to approximately 1 nmol. In the same preparation, 2 nmol of LBLOX were found. Further treatment with 100 mM unlabeled $CaCl_2$ removed 90% of the radioactivity from the lipid bodies, whereas most of the LBLOX remained bound to the lipid bodies. These data do not agree with the hypothesis that Ca^{2+} is a prerequisite for the binding of LBLOX to the surface of the lipid bodies.

A substantial modification of the LBLOX ß-barrel eliminates its ability to bind to liposomes

To further characterize the nature of the interaction between the surface of the lipid bodies and the N-terminal LBLOX ß-barrel structure, the question of whether a modified ß-barrel structure also binds to liposomes was studied. To this end, a recombinant protein (LBLOX-HA_3) was generated which, in comparison with the wild-type LBLOX, contained a triple hemagglutinin label inserted between amino acid residues 70 and 71 of LBLOX. This construction interrupts the ß-barrel by a segment of 30 amino acid residues.

The liposome experiments shown in figure 7 compile the evidence that the N-terminal LOX ß-barrel alone suffices for its binding to membranes and that destruction of the ß-barrel inactivates the transfer. Wild-type LBLOX and the ß-barrel fusion protein (GST-LOX) bind virtually quantitatively to the liposomes, while the insertion of a peptide into the barrel structure eliminates the membrane affinity.

DISCUSSION

Few isoforms of lipoxygenase bind to or integrate into the membranes of a variety of organelles. In these cases, the function of lipoxygenase can aim at a modification of the membrane in question. The expression of 15-LOX in reticulocytes reaches a maximum immediately prior to degradation of the organelle [Kühn, H. et al., J. Biol. Chem. 265, 1990: 18351-18361], and this also applies to lipid body LOX in the case

of fat-degrading cotyledons [Feussner, I. et al., Planta 198, 1998: 288-293; Matsui, K. et al., Plant. Physiol. 119, 1999: 1279-1257]. Thus, the function of LOX in several types of eukaryotic cells is a programmed organelle degradation [van Leyen, K. et al., Nature 395, 1998: 392-395]. In the case of LBLOX, both the phospholipid monolayer and most of the triacylglycerides are modified, giving rise to 13-S-hydro-peroxyoctadecadienoyl units [Feussner, I. et al. Proc. Natl. Acad. Sci. USA, 92, 1995: 11849-11853; Sturm, A. et al., Eur. J. Biochem., 150, 19985 [sic]: 461-468], which, in turn, initiates mobilization of the storage lipid. The binding of LBLOX to the target membrane may be a prerequisite for this activity.

In vitro binding experiments and cell fractionation studies which describe the situation *in vivo* were carried out in order to characterize the steps required for the directed intracellular transport of LBLOX to lipid bodies. The results of the two experimental approaches support the concept that a cytosolic pool of the primary translation product does exist and that the subsequent transfer to the lipid bodies is posttranslational. Since the target organelles are already in existence at the time of seed germination, binding of the lipid-degrading enzymes, viz. LBLOX and PLA, to the organelle depends predominantly on the transient expression of the genes encoding the particular LOX and PLA isoforms. The earlier findings of a temporary pattern of LOX expression [Feussner, I. et al., Planta 198, 1998: 288-293] and PLA expression [May. C. et al., Biochim. Biophys. Acta 1393, 1998: 267-276] in cotyledons tally with their development stage-dependent role in the degradation of lipid body structures.

It is worth pointing out that the binding observed for LBLOX is not a transient, but a stable binding to microsomes and lipid bodies. The strength of this binding was evidenced by obtaining LBLOX with lipid bodies or liposomes after the stringent separation of excess ligand and acceptor membranes by flotation. However, this binding cannot be compared with the integration of a protein comprising a long segment of hydrophobic amino acid residues, as has been shown for the oleosins, another form of lipid body proteins [Hills, M J. et al., Planta 189, 1993: 24-29]; rather, it corresponds to the behavior of peripheral membrane proteins. In the latter case, it might be useful to consider whether the binding of LBLOX takes place as the result of its interaction with an integral membrane protein as partner, for example with oleosin as anchor, or whether the function of the LBLOX ß-barrel is comparable with the function of a C2 domain which binds to the phospholipid membrane, as has been shown for synaptotagmin [Rizo, J. & Sudhof, T. C. J. Biol. Chem. 273, 1998:

15879-15882] and cytosolic phospholipase C-delta [Perisic, O. et al., J. Biol. Chem. 273, 1998: 1596-1604] or phospholipase A_2 [Xu, G. Y. et al., J. Mol. Biol. 280, 1998: 485-500]. This type of comparison with C2 domains requires an in-depth study of a potential effect of Ca^{2+} either on the membrane-binding properties of LBLOX or on the structure of LBLOX, as has been found for 5-lipoxygenase [Hammarberg, T. & Radmark, O., Biochem. 38, 1999: 4441-4447]. While experiments with lipid bodies have revealed the coating of lipid bodies with $^{45}Ca^{2+}$, no Ca^{2+} dependency of the binding of LBLOX constructs was observed. If the existence or absence of a Ca^{2+}-mediated binding is to be established unambiguously, the experimental protocols for future studies must be improved in various respects. Earlier experiments [Busch, M.B. et al., Eur. J. Cell Biol. 60, 1993: 88-100] with root tissue using energy filter electron microscopy have demonstrated that the surface of lipid bodies is covered by a zone with a high Ca^{2+} content.

We should remember that the comparison of the primary amino acid sequences and of the putative secondary structures of LOX forms shows that LBLOX is distinguished not only by the formation of a ß-barrel structure equipped with exposed glutamyl residues, but also by an N-terminal extension of 30 amino acid residues which is not present in cytosolic LOX isoforms [Höhne, M. et al., Eur. J. Biochem., 241, 1996: 6-11; Rosahl, S., Z. Naturforsch. C 51, 1996: 123-138]. It is thus probable that both types of interaction play a decisive role, namely the N-terminal extension as specific motif and the ß-barrel, which is located in the amino acid sequence which follows the N-terminal extension, as a general means for increasing membrane affinity.

Our experiment with the mutant LBLOX protein, which differs from the wild-type only by a substantial modification in the ß-barrel, i.e. by the insertion of three repeats of an HA label, demonstrated the loss of binding to lipid bodies or membranes. This means that the intact ß-barrel must be considered as being essential for the binding to lipid bodies, irrespective of whether the most N-terminal region contributes to higher selectivity or not.

Targeting signals as unfolded regions of amino acid residues have been identified when proteins are transported into mitochondria, chloroplasts or the ER. Here, where a protein binds to preexisting lipid bodies, a domain already folded into a specific shape may bring about membrane binding. A further portion which mediates the selectivity of the binding of LBLOX to lipid bodies must be postulated. If a large amount of LBLOX exists, not only

lipid bodies, but also the ER and the Golgi vesicles are covered with LBLOX. If, however, the intracellular amount of LBLOX decreases, the LBLOX molecules predominantly occupy the surface of lipid bodies.

Keys to the figures

Part A (fluorogram) shows a comparison of the migration behavior of LBLOX generated by *in vitro* translation (lane 1), LBLOX isolated from microsomes (lane 2) and LBLOX isolated from lipid bodies (lane 3). In the case of lanes 2 and 3, the corresponding cellular subfractions were prepared from cotyledons following *in vivo* protein labeling using [^{35}S]L-methionine. Parts B and C show fluorograms which show the posttranslational binding to isolated cucumber microsomes. Cucumber microsomes were incubated with the radiolabeled proteins prepared in reticulocyte lysates using either LBLOX mRNA or PLA mRNA and subsequently purified by flotation in a density gradient. The membrane-bound lipid body lipoxygenase (part B) or phospholipase (part C) which were obtained after flotation are shown on the left-hand side. Lane 1 in B and C corresponds to the upper section of the sucrose gradient, while lane 12 (of B) and lane 13 (of C) correspond to the bottom and constitute the non-membrane-bound (non-floating) proteins which remain in the position at which the incubation mixture had been applied as a layer beneath the gradient prior to centrifugation.

Fig. 2. Results of pulse-labeling proteins in cotyledons which suggest a weak binding of LBLOX to microsomes, but substantial LBLOX pool in the cytosol.

After a brief period (15 minutes), which was used to feed the radioactive amino acid as precursor to the cotyledons, a cell fractionation was carried out and two proteins were isolated from the cellular subfractions by immunoprecipitation. The radioactive proteins were isolated from the solubilized subfractions after addition of 1 µg of the corresponding cold protein (LBLOX or isocitrate lyase, ICL) as carrier by means of addition of antiserum and direct precipitation (see methods). Following electrophoresis and protein staining (lane 3: microsomes; lane 4: cytosol), the fluorogram (lanes 1 and 2) showed in part A that LOX was labeled to a high degree in the cytosol (lane 2) and considerably less in the microsomes (lane 1). Part B: the distribution of isocitrate lyase between these two fractions was determined as control. Lane 1 (microsomes) and lane 2 ("cytosol") show fluorograms, while lanes 3 and 4 represent the protein stainings. Lane 1 (and the corresponding protein staining in

lane 3) show the absence of contamination by isocitrate lyase in the microsomes. Part C: treatment of labeled microsomes (analyzed as in part A, lane 1) with proteinase K. Following proteolysis, phenylmethylsulfonyl fluoride and 1 µg of the corresponding cold protein were added, and an immunoprecipitation was carried out. Lanes 1 to 4 show fluorograms: untreated microsomes in lane 1; untreated cytosol in lane 2; treated microsomes in lane 3; treated cytosol in lane 4. Part D: localization of cucumber LBLOX at microsomes from transgenic tobacco plants. Following flotation of the membranes in a linear sucrose density gradient, subfractions were analyzed by immunoblotting. Lane 1 corresponds to the upper section of the centrifuge tube (23% sucrose), lane 3 (32% sucrose), lane 5 (39% sucrose), lane 6 (41% sucrose) and lane 7 (sample applied at 43% sucrose).

Fig. 3: *In vitro* experiments which show that radiolabeled cytosolic LBLOX binds weakly to microsome membranes (part A), but hardly at all to glyoxysomes (part B).

Part A: 1 g of cotyledons was incubated for 3 hours with 9 MBq of [^{35}S]L-methionine. Then, a 100 000 g supernatant comprising radiolabeled cytosolic LBLOX was prepared. This preparation was mixed with a homogenate prepared from untreated cotyledons. The ER/Golgi fraction was isolated by subsequent gradient centrifugation. An aliquot (1/20) of the ER/Golgi fraction (lane 1), an aliquot (1/20) of the reisolated cytosol (lane 2) and a large aliquot (1/2) of the ER/Golgi fraction (lane 3) were subjected to SDS-PAGE and fluorography. Part B: in a similar mixing experiment, isolated unlabeled glyoxysomes were incubated with the radioactive cytosol which had been prepared from in vivo-labeled cotyledons. Subsequent reisolation of the glyoxysomes and glyoxysome membranes was carried out by flotation in a sucrose gradient. To this end, the incubation mixture was adjusted to 60% (w/w) sucrose. A gradient (56 to 38% sucrose) was applied to the sample in the form of a layer. Following centrifugation at 27 000 rpm for 15 hours in a Beckman SW-28 rotor, the fractions were analyzed by SDS-PAGE and fluorography. The position of the labeled LBLOX in the fluorogram is indicated by an arrow. The lanes correspond to the following fractions (equilibrium densities in brackets): lane 4 (48% sucrose), lane 5 (48.5% sucrose), lane 6 (49% sucrose), lane 7 (50.5% sucrose), lane 8 (452.5% sucrose), lane 24 (56% sucrose), lane 25 (56.5% sucrose), lane 26 (58% sucrose), lane 27 (59% sucrose) and lane 28 (59% sucrose). The numbers of fractions 25-28 correspond to the position in the gradient at which the suspension had been applied prior to centrifugation. Lanes 4-5 comprise the

glyoxysome membranes, and lane 8 comprises the glyoxysomes, according to the protein profile.

Fig. 4. Affinity of bacterially expressed LBLOX and PLA for liposomes.

1 µg of the relevant proteins was incubated in 200 µl of buffer with an amount of liposomes corresponding to 1 mg of phosphatidylcholine. After 30 minutes, the mixture was adjusted to 42% sucrose and applied as a layer to a sucrose gradient. Flotation was carried out by centrifugation at 100 000 g for 6 hours. The proteins in the subfractions obtained from the gradient were analyzed by SDS-PAGE and immunoblots. Suitable antisera which had been raised either against LBLOX or against the patatin-like protein were used for immunolabeling. The lanes at the very right always show the bottom where the incubation mixture had been applied as a layer beneath the gradient prior to centrifugation. Thus, flotation was carried out from right to left

Fig. 5. Schematic representation of the LBLOX structure and its N-terminal section (amino acid residues 48-244) with the ß-barrel.

The LBLOX structure was calculated on the basis of the crystal structure data obtained for soybean LOX-1 [21] using the LBLOX primary sequence. The upper part of the figure shows the N-terminus, and the ß-barrel, which is exclusively composed of ß-pleated sheets, is shown slightly apart in the bottom right-hand corner. The major part of LOX, which comprises the active center at the C terminus, is dominated by α helices. The bottom part of the figure is an enlarged view of the N-terminal ß-barrel. The 40 amino acid residues of the outermost N terminus, an extension not found in other LOX structures, are not shown. The interrupted structure indicated by arrows in the bottom right-hand corner shows a location at which the amino acid sequence of the LBLOX differs considerably from that of the soybean LOX-1. The program used (Swiss 3D model, Expasy server) revealed a highly flexible loop, and thus an undefined structure, similarly as in the case of LOX-1. In LOX-1, this loop is composed of 14 amino acid residues; in the case of LBLOX, of 20 amino acid residues. The two glutamyl residues E59 and E70 are single and are only found in LBLOX and not in soybean LOX-1. This location may play a role in the Ca^{2+}-coordinated membrane association.

Fig. 6. *In vitro* bindung of the GST-LBLOX244 fusion protein to isolated lipid bodies.

The affinity-purified fusion protein GST-LBLOX244 was added to a suspension of lipid bodies in enriched cytosol. After incubation, the lipid bodies were separated from the soluble proteins by flotation. Aliquots of the two fractions and of the original crude lipid body fraction (lanes 3 and 6) were analyzed by SDS-PAGE and subsequent immunblot using anti-LBLOX antiserum. Lanes 2 and 5: lipid bodies obtained by flotation: lanes 1 and 4: reisolated proteins. The arrow pointing at the band at 51 kDa indicates the fusion protein which has bound to the lipid bodies in addition to the endogenous LBLOX. Lanes 1-3: protein staining; lanes 4-6: immunolabeling.

Fig. 7. Binding of LBLOX constructs and recombinant proteins with wild-type and modified ß-barrel to liposomes.

Following incubation of the purified recombinant proteins with liposomes, the liposomes were reisolated from the incubation mixture by flotation in a sucrose gradient. The behavior of wild-type lipid-body LOX (LBLOX) was compared with that of a fragment comprising the N-terminal half of LBLOX (LBLOX-Δ504), with the fusion protein with the ß-barrel (GST-LOX) and with an LBLOX whose N-terminal ß-barrel had been modified substantially (LBLOX-HA_3). The figure shows the analysis of gradient fractions by SDS-PAGE and immunolabeling using anti-LBLOX antiserum.

Abbreviations. Hemagglutinin, HA; glutathione-S transferase, GST; lipid body lipoxygenase, LBLOX; lipoxygenase, LOX; phospholipase A_2, PLA.

Enzymes. Glutathione-S transferase (EC 2.5.1.18); isocitrate lyase (EC 4.1.3.1); lipoxygenase (EC 1.13.11.12); phospholipase A_2 (EC 3.1.4.3).